UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/748,542	12/26/2000	Nobuyasu Itoh	6169-169	3020

40987 7590 05/28/2004

AKERMAN SENTERFITT
P. O. BOX 3188
WEST PALM BEACH, FL 33402-3188

EXAMINER	
LEWIS, MICHAEL A	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 05/28/2004

10

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 09/748,542	Applicant(s) ITOH ET AL.	
	Examiner Lewis A Michael	Art Unit 2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 - 48* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 - 15, 23 - 25, 35 - 37 & 47 - 48* is/are rejected.

7)☒ Claim(s) *16 - 22, 26 - 34 & 38 - 46* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *05*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Allowable Subject Matter

1. Claims 16 – 22, 26 – 34 & 38 – 46 objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

2. The following is a statement of reasons for the indication of allowable subject matter:

Regarding claims 16 - 22, 26 - 34 & 38 - 46, the claim limitation pertaining to a third or fourth or fifth probability calculator, wherein said third or fourth or fifth probability calculator calculates said probability based on said first and/or second language model, if said conditional words have been judged as containing only non-disfluency words by said second judging processor is not taught in prior art.

In addition, the claim limitation pertaining to a third judging processor, wherein said third judging processor judges whether a word immediately preceding said object word is a disfluency word; and a fourth or fifth probability calculator, wherein said fourth or fifth probability calculator calculates said probability based on said first and/or said second language models, if said preceding word has been judged a disfluency word by said third judging processor is not taught in prior art.

The combination of Padmanabhan et al., Tang et al. and Stolcke et al. teach a second processor with a first language model and a second probability calculator to that deal with words that have been judged non-disfluent(Fig 2 (3,5, 8)). However, the combination of Padmanabhan et al., Tang et al. and Stolcke et al. do not teach a third or fourth or fifth probability calculator, wherein said third or fourth or fifth probability calculator calculates said probability based on said first and/or second language model, if said conditional words have been judged as containing only non-disfluency words by said second judging processor.

The combination of Padmanabhan et al., Tang et al. and Stolcke et al. teach the use of a judging processor and a fourth and fifth probability function for determining the location and type of disfluenct word. The combination of Padmanabhan et al., Tang et al. and Stolcke et al. do not teach a third judging processor that judges whether a word immediately preceding said object word is a disfluency word; and a fourth or fifth probability calculator, wherein said fourth or fifth probability calculator calculates said probability based on said first and/or said second language models, if said preceding word has been judged a disfluency word by said third judging processor.

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1, 2, 5, 6, 7, 8, 9, 10, 11, 12,47 & 48 are rejected under 35 U.S.C. 103(a) as being unpatentable over Padmanabhan et al. (US6385579) in view of Tang et al.(US 6718303) and further in view of Stolcke et al.(IEEE 0-7803-3192-3/96).

In regards to claims 1, 5, 7, 9,11, 47 & 48, Padmanabhan et al. disclose an apparatus/method with storage medium for speech recognition, comprising: an acoustic processor, wherein said acoustic processor converts analog speech input signals into digital signals (Fig 1(10 & 12); Fig 2(40))); a first storage structure, wherein said first storage structure stores an acoustic model which has learned voice characteristics (Fig 1(20 & 60)); a probability regarding said digital signals using said acoustic model and said dictionary to recognize words showing the highest probability of representing said input signals(Fig 1(14);equation 5,7 & 8).

In addition, Padmanabhan et al. disclose a second storage structure

[Compound Word Formation Module], wherein said second storage structure stores a dictionary containing a first language model (Fig 1(22)).

Padmanabhan et al. do not disclose explicitly the storing of a dictionary containing a first language model that has been trained regarding disfluency words and non-disfluency words, and a second language model which has been trained regarding non-disfluency words and trained to ignore disfluency words.

Tang et al. teaches a second storage structure, wherein said second storage structure stores a dictionary containing a first language model which has been trained related to general words within the language and *a second language* model which contains speech with all the pseudo punctuation marks associated with many disfluencies such as silence, lip smacking etc. (Col 3, Line 54 – 59;Col 5, Lines 20 – 21; Col5, Line 40 - 45). Tang does not explicitly teach a language model trained regarding disfluency words and non-disfluency words, and another language model that has been trained regarding non-disfluency words and trained to ignore disfluency words. However, Stolcke et al. teach a language model trained regarding disfluency words and non-disfluency words, and another language model *[baseline model]* that has been trained regarding non-disfluency words and trained to ignore disfluency words (page 406, Section 3). It is beneficial to recognize speech disfluencies using a

language model that contains word prediction capabilities for more accurate results in spontaneous speech recognition systems.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Tang et al. with use of language models trained to recognize nondisfluent and disfluent words as taught by Stolcke since it is beneficial to recognize speech disfluencies using a language model that contains word prediction capabilities for more accurate results in spontaneous speech recognition systems.

The combination of Tang et al. and Stolcke et al. modifies Padmanabhan et al. to teach the storing of a dictionary containing a first language model that has been trained regarding disfluency words and non-disfluency words, and a second language model which has been trained regarding non-disfluency words and trained to ignore disfluency words.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Padmanabhan et al. with the use of language models trained to recognize nondisfluent and disfluent words as taught by the combination of Tang et al. and Stolcke et al. since the specific training would have created a more accurate spontaneous speech recognition system.

Regarding claims 2, 6, 8, 10 & 12, Padmanabhan et al. do not disclose that the first and second language models are N-gram models. However, the combination of Tang et al. and Stolcke et al. teach said first and second language models are N-gram models (Stolke (Abstract)). Dynamic programming to compute the probability of a word sequences/prediction as it relates to disfluent events uses N-Gram models.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Padmanabhan et al. with the use N-gram models as taught by the combination of Tang et al. and Stolcke et al. it would have been beneficial to N-Gram models for dynamic programming to compute the probability of a word sequences/prediction.

3. Claims 3 & 4 are rejected under 35 U.S.C. 103(a) as being unpatentable over Padmanabhan et al. (US6385579) in view of Tang et al. (US 6718303) and further in view of Stolcke et al.(IEEE 0-7803-3192-3/96) and further in view of Bellegarda (US6374217).

Regarding claim 3, the modified Padmanabhan et al. do not disclose a

display apparatus for displaying results of said recognition. However, Bellegarda teaches the use of monitor to display the output of the said recognition (Fig 2(221)).

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Padmanabhan et al. with the use of a display as presented by Bellegarda since it would have been beneficial to the user to see an output of the recognized result.

Regarding claim 4, Padmanabhan et al. do not disclose that the first and second language models are N-gram models. However, the combination of Tang et al. and Stolcke et al. teaches said first and second language models are N-gram models (Stolcke (Abstract)). Dynamic programming to compute the probability of a word sequences/prediction as it relates to disfluent events uses N-Gram models.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Padmanabhan et al. with the use N-gram models as taught by the combination of Tang et al. and Stolcke et al. it would have been beneficial to N-Gram models for dynamic programming to compute the probability of a word sequences/prediction.

4. Claims 13, 14, 15, 23, 35, 24, 25, 36 & 37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Tang et al. (US 6718303) and further in view of Stolcke et al. (IEEE 0-7803-3192-3/96).

Regarding claim 13, Tang et al. disclose an apparatus for recognizing speech from texts comprising disfluency words and non-disfluency words, said apparatus comprising: a first judging processor, wherein said first judging processor judges whether words inputted as an object of recognition are non-disfluency words *[Tang et al. describe an acoustic model which together with the probability determines the pronunciation of each word and the pseudo noise (which includes some types of disfluencies)] (Page 3, Paragraph 0031 – 0032)*; a second judging processor, wherein said second judging processor judges whether said inputted words constituting a condition necessary for recognizing said inputted words consist of only non-disfluency words, if said inputted words have been judged to be non-disfluency words by said first judging processor *[Tang et al. describes a word matching means(processor) and context generating means which is responsible for recognizing the input word only]*(Fig 3(3); Page 3, Paragraph 9931); and a first probability calculator, wherein said first probability calculator calculates a probability, if said conditional words have been judged as containing non-disfluency words and disfluency words by said second judging processor, by using a dictionary containing a first language model which has been trained

regarding disfluency words and non-disfluency words, and a second language model which has been trained regarding non-disfluency words and trained to ignore disfluency words so as to recognize words showing the highest probability of representing said inputted words *[Tang et al. in (Fig 2 (5, 8) ; Page 3, Paragraph 0033) describes a word probability calculator and a language model that has been trained to recognize/distinguish between words and pseudo noises(disfluencies)].*

Tang et al. teach the use of pseudo noises which include silence, pause , lip smacking etc. however, Tang et al. do not describe dis-fluencies such as repetitions, deletions and other disfluent words such as "umhs", "uh", etc. However, Stolcke et al. teach a language model trained regarding disfluency words and non-disfluency words, and another language model *[baseline or reference model]* that has been trained regarding non-disfluency words and trained to ignore disfluency words (page 406, Section 3). It is beneficial to recognize speech all disfluencies using a language model that contains word prediction capabilities for more accurate results in spontaneous speech recognition systems.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Tang et al. with use of language models trained to recognize a broader array of non-disfluent words as taught by Stolcke

since a broader array of disfluencies using a language model would have resulted in the creation of a more accurate spontaneous speech recognition system.

Regarding claim 14, Tang et al. disclose an apparatus for speech recognition further comprising: a second probability calculator, wherein said second probability calculator calculates said probability based on said first language model, if said object words have been judged as **not being** non-disfluency words by said first judging processor (Fig 2(10)).

Regarding claim 15, Tang et al. disclose an apparatus for speech recognition further comprising: a third probability calculator, wherein said third probability calculator calculates probability based on said second language model, if said conditional words have been judged as containing only non-disfluency words by said second judging processor (Fig 2 (13)).

Regarding claims 23 & 35, Tang et al. disclose amethod for recognizing speech from texts comprising disfluency words and non-disfluency words, comprising the steps of: (a) judging whether words inputted as an object of recognition are non-disfluency words *[Tang et al. describe an acoustic model which together with the probability determines the pronunciation of each word and the pseudo noise (which includes some types of*

disfluencies)] (Col 3, Paragraph 0031 – 0032); (b) judging further whether said words constituting a condition necessary for recognizing said input words consist only of non-disfluency words, if said inputted words have been judged to be non-disfluency words in said step (a) *[Tang et al. describes a word matching means processor and context generating means which is responsible for recognizing the input word only]*(Fig 3(3); Page 3, Paragraph 9931)); and (c) calculating a probability, if said conditional words have been judged as comprising non-disfluency words and disfluency words in said step (b), by using a dictionary containing a first language model which has been trained regarding disfluency words and non-disfluency words, and a second language model which has been trained regarding non-disfluency words and trained to ignore disfluency words so as to recognize words showing the highest probability of representing said input words *[Tang et al. in (Fig 2 (5, 8); Page 3, Paragraph 0033) describes a word probability calculator and a language model that has been trained to recognize words and pseudo noises (dis-fluencies)]*.

Tang et al. teach the use of pseudo noises which includes silence, pause , lip smacking etc. however, Tang et al. do not describe disfluencies such as repetitions, deletions and other disfluent words such as “umhs”, “uh”, etc. However, Stolcke et al. teach a language model trained regarding

disfluency words and non-disfluency words, and another language model *[baseline model]* that has been trained regarding non-disfluency words and trained to ignore disfluency words (page 406, Section 3). It is beneficial to recognize speech disfluencies using a language model that contains word prediction capabilities for more accurate results in spontaneous speech recognition systems.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Tang et al. with use of language models trained to recognize a broader array of non-disfluent words as taught by Stolcke since a broader array of disfluencies using a language model would have resulted in the creation of a more accurate spontaneous speech recognition system.

Regarding claim 24 & 36, Tang et al. disclose The method for speech recognition further comprising the step of: calculating said probability based on said first language model, if said object words have been judged as not being non-disfluency words in said step (a) (Fig 2(10)).

Regarding claim 25 & 37, Tang et al. disclose the method for speech recognition further comprising the step of: calculating said probability based on said second language model, if said conditional words have

been judged as consisting only of non-disfluency words in said step (b)(Fig 2 (13)).

Conclusion

1. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Bokser	U.S. Patent (5261009)
Dahan et al.	U.S. Patent (6018708)
Soufflet et al.	U.S. Patent Application (20030093272)
Deligne et al.	U.S. Patent (6314399)
Sukkar	U.S. Patent (6292778)
Modi et al.	U.S. Patent (6125345)
Sherwood et al.	U.S. Patent (6212498)
Chou et al.	U.S. Patent (5797123)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael A Lewis whose telephone number is 703 305-8730. The examiner can normally be reached on Monday through Friday, 8:30 am – 5 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on (703)305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Lewis A Michael
Examiner
Art Unit 2655

Mal

5/1/2004

DORIS H. TO
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600